

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 09/360,685D  
Source: 1FW/b  
Date Processed by STIC: 6/26/06

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 06/26/2006

PATENT APPLICATION: US/09/360,685D

TIME: 11:30:33

Input Set : E:\CHIR0157 (updated 06-06-06).ST25.txt

Output Set : N:\CRF4\06262006\I360685D.raw

```

3 <110> APPLICANT: Covacci, Antonello
4     Bugnoli, Massimo
5     Telford, John
6     Macchia, Giovanni
7     Rappuoli, Rino
9 <120> TITLE OF INVENTION: Helicobacter Pylori CAI Antigen Polynucleotides
11 <130> FILE REFERENCE: CHIR0157
13 <140> CURRENT APPLICATION NUMBER: 09/360,685D
14 <141> CURRENT FILING DATE: 1999-07-26
16 <150> PRIOR APPLICATION NUMBER: 08/471,491
17 <151> PRIOR FILING DATE: 1995-06-06
19 <150> PRIOR APPLICATION NUMBER: 08/256,848
20 <151> PRIOR FILING DATE: 1994-10-21
22 <160> NUMBER OF SEQ ID NOS: 30
24 <170> SOFTWARE: PatentIn version 3.3
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 27
28 <212> TYPE: DNA
29 <213> ORGANISM: Artificial Sequence
31 <220> FEATURE:
32 <223> OTHER INFORMATION: primer oligonucleotide
34 <400> SEQUENCE: 1
35 gcaagcttat cgatgtcgac tcgagct
38 <210> SEQ ID NO: 2
39 <211> LENGTH: 3960
40 <212> TYPE: DNA
41 <213> ORGANISM: Helicobacter pylori
43 <400> SEQUENCE: 2
44 aaaaagaaag gaagaaaatg gaaatacaac aaacacaccg caaaatcaat cgccctctgg      60
46 tttctctcgc tttagtagga gcattagtcg gcatcacacc gcaacaaaagt catgccgcct      120
48 ttttcacaac cgtgatcatt ccagccattg ttgggggtat cgctacaggc accgctgtag      180
50 gaacgggtctc agggcttctt agctgggggc tcaaacaagc cgaagaagcc aataaaaccc      240
52 cagataaacc cgataaagtt tggcgcatct aagcaggaaa aggcctttaa gaattcccta      300
54 acaaggaata cgacttatac agatcccttt tatccagtaa gattgatgga ggttgggatt      360
56 gggggaatgc cgctaggcat tattgggtca aaggcgggca acagaataag cttgaagtgg      420
58 atatgaaaga cgctgtaggg acttatacct tatcagggct tagaaacttt actggtgggg      480
60 atttagatgt caatatgcaa aaagccactt tacgcttggg ccaattcaat ggcaattctt      540
62 ttacaagcta taaggatagt gctgatcgca ccacgagagt ggatttcaac gctaaaaata      600
64 tctcaattga taattttgta gaaatcaaca atcgtgtggg ttctggagcc gggaggaaaag      660
66 ccagctctac ggttttgact ttgcaagctt cagaagggat cactagcgat aaaaacgctg      720
68 aaatttctct ttatgatggt gccacgctca atttggtctc aagcagcggt aaattaatgg      780
70 gtaatgtgtg gatgggccgt ttgcaatacg tgggagcgta tttggccctt tcatacagca      840
72 cgataaacac ttcaaaaagta acaggggaag tgaattttaa ccacctcact gttggcgata      900

```

27

## RAW SEQUENCE LISTING

DATE: 06/26/2006

PATENT APPLICATION: US/09/360,685D

TIME: 11:30:33

Input Set : E:\CHIR0157 (updated 06-06-06).ST25.txt

Output Set: N:\CRF4\06262006\I360685D.raw

74	aaaacgccgc	tcaagcgggc	attatcgcta	ataaaaagac	taatattggc	acactggatt	960
76	tgtggcaaa	cgccgggta	aacattatcg	ctcctccaga	aggtggctat	aaggataaac	1020
78	ccaataatac	cccttctcaa	agtggtgcta	aaaacgacaa	aaatgaaagc	gctaaaaacg	1080
80	acaaacaaga	gagcagtcaa	aataatagta	acactcaggt	cattaaccca	cccaatagt	1140
82	cgcaaaaaac	agaagttcaa	cccacgcaag	tcattgatgg	gccttttgcg	ggcggcaaa	1200
84	acacggttgt	caatatcaac	cgcacgaaca	ctaacgctga	tggcacgatt	agagtgggag	1260
86	ggttttaaagc	ttctcttacc	accaatgcgg	ctcatttgc	tatcggaaca	ggcgggtgtca	1320
88	atctgtccaa	tcaagcgagc	gggcgctctc	ttatagtggg	aaatctaact	gggaatatca	1380
90	ccgttgatgg	gcctttaaga	gtgaataatc	aagtgggtgg	ctatgctttg	gcaggatcaa	1440
92	gcgcgaattt	tgagtttaag	gctggtacgg	ataccaaaaa	cggcacagcc	acttttaata	1500
94	acgatattag	tctgggaaga	tttgtgaatt	taaaggtgga	tgctcataca	gctaatttta	1560
96	aaggtattga	tacgggtaat	ggtggtttca	acaccttaga	ttttagtggc	gttacagaca	1620
98	aagtcaatat	caacaagctc	attacggctt	ccactaatgt	ggccgttaaa	aacttcaaca	1680
100	ttaatgaatt	gattgttaaa	accaatggga	taagtgtggg	ggaatatact	catttttagcg	1740
102	aagatatagg	cagtcaatcg	cgcacgaaca	ccgtgcgttt	ggaaactggc	actaggtcac	1800
104	ttttctctgg	gggtgttaaa	tttaaggtg	gcgaaaaatt	ggttatagat	gagttttact	1860
106	atagcccttg	gaattatttt	gacgctagaa	atattaaaaa	tgttgaaatc	accaataaac	1920
108	ttgcttttgg	acctcaagga	agtccttggg	gcacatcaaa	acttatgttc	aataatctaa	1980
110	ccctagggtca	aaatgcggtc	atggattata	gccaattttc	aaatttaacc	attcaagggg	2040
112	atttctacaa	caatcaaggc	actatcaact	atctggtccg	aggtgggaaa	gtggcaacct	2100
114	taagcgtagg	caatgcagca	gctatgatgt	ttaataatga	tatagacagc	gcgaccggat	2160
116	tttacaacc	gctcatcaag	attaacagcg	ctcaagatct	cattaaaaat	acagaacatg	2220
118	ttttattgaa	agcgaaaatc	attggttatg	gtaatgtttc	tacaggtacc	aatggcatta	2280
120	gtaatgttaa	tctagaagag	caattcaaag	agcgccctagc	cctttataac	aacaataacc	2340
122	gcatggatac	ttgtgtggtg	cgaaatactg	atgacattaa	agcatgcggg	atggctatcg	2400
124	gcgatcaaag	catgggtgaac	aaccctgaca	attacaagta	tcttatcggt	aaggcatgga	2460
126	aaaatatagg	gatcagcaaa	acagctaatt	gctctaaaat	ttcgggtgat	tatttaggca	2520
128	attctacgcc	tactgagaat	ggtggcaata	ccacaaat	accacaaaac	accactagca	2580
130	atgcacgttc	tgccaacaac	gcccttgac	aaaacgctcc	tttcgctcaa	cctagtgtca	2640
132	ctcctaattt	agtcgctatc	aatcagcatg	attttggcac	tattgaaagc	gtgtttgaat	2700
134	tggttaaccg	ctctaaagat	attgacacgc	tttatgctaa	ctcaggcgct	caaggcaggg	2760
136	atctcttaca	aaccttattg	attgatagcc	atgatgcggg	ttatgccaga	aaaatgattg	2820
138	atgctacaag	cgctaataag	atcaccaagc	aattgaatac	ggccactacc	actttaaaca	2880
140	acatagccag	tttagagcat	aaaaccagcg	gcttacaac	tttgagcttg	agtaatgcga	2940
142	tgatttttaa	ttctcgttta	gtcaatctct	ccaggagaca	caccaacct	attgactcgt	3000
144	tcgccaacg	cttacaagct	ttaaaagacc	aaaaattcgc	ttctttagaa	agcgcgag	3060
146	aagtgttgta	tcaatttgcc	cctaaatatg	aaaaacctac	caatgtttgg	gctaacgcta	3120
148	ttgggggaac	gagcttgaat	aatggctcta	acgcttcatt	gtatggcaca	agcgcgggcg	3180
150	tagacgctta	ccttaacggg	caagtgggaag	ccattgtggg	cggttttgga	agctatgggt	3240
152	atagctcttt	taataatcgt	gcgaactccc	tttaactctgg	ggccaataac	actaattttg	3300
154	gcgtgtatag	ccgtattttt	gccaaccagc	atgaatttga	ctttgaagct	caaggggcac	3360
156	tagggagcga	tcaatcaagc	ttgaatttca	aaagcgctct	attacaagat	ttgaatcaaa	3420
158	gctatcatta	cttagcctat	agcgctgcaa	caagagcgag	ctatggttat	gacttcgcgt	3480
160	tttttaggaa	cgcttttagt	ttaaaaccaa	gcgtgggtgt	gagctataac	catttagggt	3540
162	caaccaactt	taaaagcaac	agcaccaatc	aagtggcttt	gaaaaatggc	tctagcagtc	3600
164	agcattttatt	caacgctagc	gctaattgtg	aagcgcgcta	ttattatggg	gacacttcat	3660
166	acttctacat	gaatgctgga	gttttacaag	agttcgctca	tgttggctct	aataacgccg	3720
168	cgtcttttaa	caccttttaa	gtgaatgccg	ctcgcaaccc	tttaaatacc	catgccagag	3780
170	tgatgatggg	tggggaatta	aaattagcta	aagaagtgtt	tttgaatttg	ggcgttgttt	3840

## RAW SEQUENCE LISTING

DATE: 06/26/2006

PATENT APPLICATION: US/09/360,685D

TIME: 11:30:33

Input Set : E:\CHIR0157 (updated 06-06-06).ST25.txt

Output Set: N:\CRF4\06262006\I360685D.raw

```

172 atttgacaaa ttgatttcc aatataggcc atttcgcttc caatttagga atgaggtata 3900
174 gtttctaaat accgctctta aaccatgct caaagcatgg gtttgaaatc ttacaaaaca 3960
177 <210> SEQ ID NO: 3
178 <211> LENGTH: 1296
179 <212> TYPE: PRT
180 <213> ORGANISM: Helicobacter pylori
182 <400> SEQUENCE: 3
184 Met Glu Ile Gln Gln Thr His Arg Lys Ile Asn Arg Pro Leu Val Ser
185 1 5 10 15
188 Leu Ala Leu Val Gly Ala Leu Val Ser Ile Thr Pro Gln Gln Ser His
189 20 25 30
192 Ala Ala Phe Phe Thr Thr Val Ile Ile Pro Ala Ile Val Gly Gly Ile
193 35 40 45
196 Ala Thr Gly Thr Ala Val Gly Thr Val Ser Gly Leu Leu Ser Trp Gly
197 50 55 60
200 Leu Lys Gln Ala Glu Glu Ala Asn Lys Thr Pro Asp Lys Pro Asp Lys
201 65 70 75 80
204 Val Trp Arg Ile Gln Ala Gly Lys Gly Phe Asn Glu Phe Pro Asn Lys
205 85 90 95
208 Glu Tyr Asp Leu Tyr Arg Ser Leu Leu Ser Ser Lys Ile Asp Gly Gly
209 100 105 110
212 Trp Asp Trp Gly Asn Ala Ala Arg His Tyr Trp Val Lys Gly Gly Gln
213 115 120 125
216 Gln Asn Lys Leu Glu Val Asp Met Lys Asp Ala Val Gly Thr Tyr Thr
217 130 135 140
220 Leu Ser Gly Leu Arg Asn Phe Thr Gly Gly Asp Leu Asp Val Asn Met
221 145 150 155 160
224 Gln Lys Ala Thr Leu Arg Leu Gly Gln Phe Asn Gly Asn Ser Phe Thr
225 165 170 175
228 Ser Tyr Lys Asp Ser Ala Asp Arg Thr Thr Arg Val Asp Phe Asn Ala
229 180 185 190
232 Lys Asn Ile Ser Ile Asp Asn Phe Val Glu Ile Asn Asn Arg Val Gly
233 195 200 205
236 Ser Gly Ala Gly Arg Lys Ala Ser Ser Thr Val Leu Thr Leu Gln Ala
237 210 215 220
240 Ser Glu Gly Ile Thr Ser Asp Lys Asn Ala Glu Ile Ser Leu Tyr Asp
241 225 230 235 240
244 Gly Ala Thr Leu Asn Leu Ala Ser Ser Ser Val Lys Leu Met Gly Asn
245 245 250 255
248 Val Trp Met Gly Arg Leu Gln Tyr Val Gly Ala Tyr Leu Ala Pro Ser
249 260 265 270
252 Tyr Ser Thr Ile Asn Thr Ser Lys Val Thr Gly Glu Val Asn Phe Asn
253 275 280 285
256 His Leu Thr Val Gly Asp Lys Asn Ala Ala Gln Ala Gly Ile Ile Ala
257 290 295 300
260 Asn Lys Lys Thr Asn Ile Gly Thr Leu Asp Leu Trp Gln Ser Ala Gly
261 305 310 315 320
264 Leu Asn Ile Ile Ala Pro Pro Glu Gly Gly Tyr Lys Asp Lys Pro Asn
265 325 330 335

```

## RAW SEQUENCE LISTING

DATE: 06/26/2006

PATENT APPLICATION: US/09/360,685D

TIME: 11:30:33

Input Set : E:\CHIR0157 (updated 06-06-06).ST25.txt

Output Set: N:\CRF4\06262006\I360685D.raw

```

268 Asn Thr Pro Ser Gln Ser Gly Ala Lys Asn Asp Lys Asn Glu Ser Ala
269          340          345          350
272 Lys Asn Asp Lys Gln Glu Ser Ser Gln Asn Asn Ser Asn Thr Gln Val
273          355          360          365
276 Ile Asn Pro Pro Asn Ser Ala Gln Lys Thr Glu Val Gln Pro Thr Gln
277          370          375          380
280 Val Ile Asp Gly Pro Phe Ala Gly Gly Lys Asp Thr Val Val Asn Ile
281 385          390          395          400
284 Asn Arg Ile Asn Thr Asn Ala Asp Gly Thr Ile Arg Val Gly Gly Phe
285          405          410          415
288 Lys Ala Ser Leu Thr Thr Asn Ala Ala His Leu His Ile Gly Lys Gly
289          420          425          430
292 Gly Val Asn Leu Ser Asn Gln Ala Ser Gly Arg Ser Leu Ile Val Glu
293          435          440          445
296 Asn Leu Thr Gly Asn Ile Thr Val Asp Gly Pro Leu Arg Val Asn Asn
297          450          455          460
300 Gln Val Gly Gly Tyr Ala Leu Ala Gly Ser Ser Ala Asn Phe Glu Phe
301 465          470          475          480
304 Lys Ala Gly Thr Asp Thr Lys Asn Gly Thr Ala Thr Phe Asn Asn Asp
305          485          490          495
308 Ile Ser Leu Gly Arg Phe Val Asn Leu Lys Val Asp Ala His Thr Ala
309          500          505          510
312 Asn Phe Lys Gly Ile Asp Thr Gly Asn Gly Gly Phe Asn Thr Leu Asp
313          515          520          525
316 Phe Ser Gly Val Thr Asp Lys Val Asn Ile Asn Lys Leu Ile Thr Ala
317          530          535          540
320 Ser Thr Asn Val Ala Val Lys Asn Phe Asn Ile Asn Glu Leu Ile Val
321 545          550          555          560
324 Lys Thr Asn Gly Ile Ser Val Gly Glu Tyr Thr His Phe Ser Glu Asp
325          565          570          575
328 Ile Gly Ser Gln Ser Arg Ile Asn Thr Val Arg Leu Glu Thr Gly Thr
329          580          585          590
332 Arg Ser Leu Phe Ser Gly Gly Val Lys Phe Lys Gly Gly Glu Lys Leu
333          595          600          605
336 Val Ile Asp Glu Phe Tyr Tyr Ser Pro Trp Asn Tyr Phe Asp Ala Arg
337          610          615          620
340 Asn Ile Lys Asn Val Glu Ile Thr Asn Lys Leu Ala Phe Gly Pro Gln
341 625          630          635          640
344 Gly Ser Pro Trp Gly Thr Ser Lys Leu Met Phe Asn Asn Leu Thr Leu
345          645          650          655
348 Gly Gln Asn Ala Val Met Asp Tyr Ser Gln Phe Ser Asn Leu Thr Ile
349          660          665          670
352 Gln Gly Asp Phe Ile Asn Asn Gln Gly Thr Ile Asn Tyr Leu Val Arg
353          675          680          685
356 Gly Gly Lys Val Ala Thr Leu Ser Val Gly Asn Ala Ala Ala Met Met
357          690          695          700
360 Phe Asn Asn Asp Ile Asp Ser Ala Thr Gly Phe Tyr Lys Pro Leu Ile
361 705          710          715          720
364 Lys Ile Asn Ser Ala Gln Asp Leu Ile Lys Asn Thr Glu His Val Leu

```

## RAW SEQUENCE LISTING

DATE: 06/26/2006

PATENT APPLICATION: US/09/360,685D

TIME: 11:30:33

Input Set : E:\CHIR0157 (updated 06-06-06).ST25.txt

Output Set: N:\CRF4\06262006\I360685D.raw

```

365          725          730          735
368 Leu Lys Ala Lys Ile Ile Gly Tyr Gly Asn Val Ser Thr Gly Thr Asn
369          740          745          750
372 Gly Ile Ser Asn Val Asn Leu Glu Glu Gln Phe Lys Glu Arg Leu Ala
373          755          760          765
376 Leu Tyr Asn Asn Asn Asn Arg Met Asp Thr Cys Val Val Arg Asn Thr
377          770          775          780
380 Asp Asp Ile Lys Ala Cys Gly Met Ala Ile Gly Asp Gln Ser Met Val
381 785          790          795          800
384 Asn Asn Pro Asp Asn Tyr Lys Tyr Leu Ile Gly Lys Ala Trp Lys Asn
385          805          810          815
388 Ile Gly Ile Ser Lys Thr Ala Asn Gly Ser Lys Ile Ser Val Tyr Tyr
389          820          825          830
392 Leu Gly Asn Ser Thr Pro Thr Glu Asn Gly Gly Asn Thr Thr Asn Leu
393          835          840          845
396 Pro Thr Asn Thr Thr Ser Asn Ala Arg Ser Ala Asn Asn Ala Leu Ala
397          850          855          860
400 Gln Asn Ala Pro Phe Ala Gln Pro Ser Ala Thr Pro Asn Leu Val Ala
401 865          870          875          880
404 Ile Asn Gln His Asp Phe Gly Thr Ile Glu Ser Val Phe Glu Leu Ala
405          885          890          895
408 Asn Arg Ser Lys Asp Ile Asp Thr Leu Tyr Ala Asn Ser Gly Ala Gln
409          900          905          910
412 Gly Arg Asp Leu Leu Gln Thr Leu Leu Ile Asp Ser His Asp Ala Gly
413          915          920          925
416 Tyr Ala Arg Lys Met Ile Asp Ala Thr Ser Ala Asn Glu Ile Thr Lys
417          930          935          940
420 Gln Leu Asn Thr Ala Thr Thr Thr Leu Asn Asn Ile Ala Ser Leu Glu
421 945          950          955          960
424 His Lys Thr Ser Gly Leu Gln Thr Leu Ser Leu Ser Asn Ala Met Ile
425          965          970          975
428 Leu Asn Ser Arg Leu Val Asn Leu Ser Arg Arg His Thr Asn His Ile
429          980          985          990
432 Asp Ser Phe Ala Lys Arg Leu Gln Ala Leu Lys Asp Gln Lys Phe Ala
433          995          1000          1005
436 Ser Leu Glu Ser Ala Ala Glu Val Leu Tyr Gln Phe Ala Pro Lys
437          1010          1015          1020
440 Tyr Glu Lys Pro Thr Asn Val Trp Ala Asn Ala Ile Gly Gly Thr
441          1025          1030          1035
444 Ser Leu Asn Asn Gly Ser Asn Ala Ser Leu Tyr Gly Thr Ser Ala
445          1040          1045          1050
448 Gly Val Asp Ala Tyr Leu Asn Gly Gln Val Glu Ala Ile Val Gly
449          1055          1060          1065
452 Gly Phe Gly Ser Tyr Gly Tyr Ser Ser Phe Asn Asn Arg Ala Asn
453          1070          1075          1080
456 Ser Leu Asn Ser Gly Ala Asn Asn Thr Asn Phe Gly Val Tyr Ser
457          1085          1090          1095
460 Arg Ile Phe Ala Asn Gln His Glu Phe Asp Phe Glu Ala Gln Gly
461          1100          1105          1110

```

**VERIFICATION SUMMARY**

DATE: 06/26/2006

PATENT APPLICATION: US/09/360,685D

TIME: 11:30:34

Input Set : E:\CHIR0157 (updated 06-06-06).ST25.txt

Output Set: N:\CRF4\06262006\I360685D.raw